Virtualization in ATLAS

Yushu Yao, Lawrence Berkeley National Lab

March 08/09 OSG All Hands Meeting

Advantages

- Most Immediate: Consolidation of Resource (Maybe)
- Independent Environment, one physical cluster can accommodate more than one experiments easier
- Help simplify deployment, both Workstations and Clusters
- Run our software 10 years later

The ATLAS Tier3 VM Workgroup

- March - May 2010
- Find and Document existing work in ATLAS
- Try to provide use case analysis and best practice guidelines
- Observations and recommendations
- This surely will not cover all the existing work inside ATLAS.
- Please feel free to contact me for any suggestions/contributions

Existing Work in ATLAS

- CernVM
  - As a Workstation for Development/Analysis
  - As a batch node
- Server Consolidation, putting many tier3 servers in the same physical node
  - VMs for XRood Redirector, Proxy, Proof Master, Condor Head, LDAP
- Performance Tests (Network, File System, Compute, mixed results)
- Cluster Automation (Puppet, CRV)

Tasks To Do

- To get a really working VM based Tier3, we need:
  - Test network file system and find a good one
  - Test Hypervisors and give suggestions
  - Test performance with multiple/parallel jobs, and scale it up
  - Test it on the Cloud (e.g. Nimbus, EC2, Eucalyptus, etc)
  - Auto-Deployment method to minimize human effort
- Some of the latest CPU features, e.g. Hyper-threading, will it be useable in VMs

Cloud

- Will a Tier3-Cloud (in addition to t2g, t2w, etc) be feasible?
  - Calculation suggests currently it is not economical to run Tier3 on EC2, unless they hold ATLAS data for free.
- A “science cloud” will be more pratical
- Ongoing research projects like Nimbus or Magellan will give valuable info

Cloud vs. Local, costs

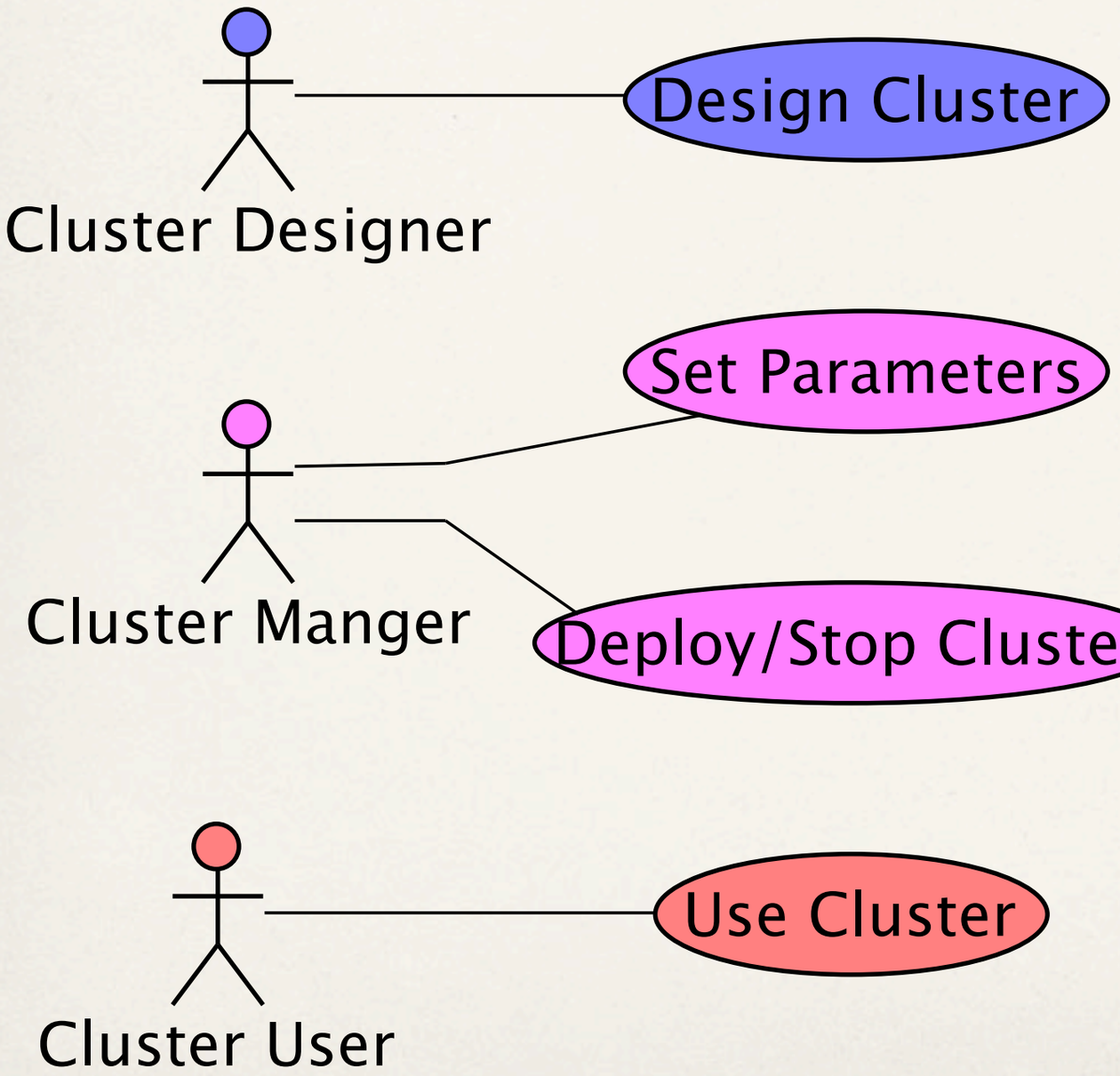
- 3yr cost for: 80 cores, 2GB per Core, 25TB Storage

	BuyReal	EC2
Initial Cost	50k (10by8Core machine+DataServer)	28k(to buy 3yr reservation)
Power (AlwaysOn/HalfTime)	21k/11k (250w by 10 always-on, 20c/kwh)	0
Storage, etc	Site Dependent	0
Man Power	0.2FTE	0.2FTE, might be less
Usage (AlwaysOn/HalfTime)	0	60k/30k
Data In/Out/Storage	0	0/0/160k (Keep 25TB storage)
Total 3yr Cost: (AlwaysOn/HalfTime)	~70k/61k	~88k/58k compute +160k Data

Cluster Automation

- Need a small tool that:
  - Automatically deploy a cluster on a physical farm or in the Cloud
  - Monitor the Cluster for its functionality
  - Scale up or down according to demand.

Cluster Automation

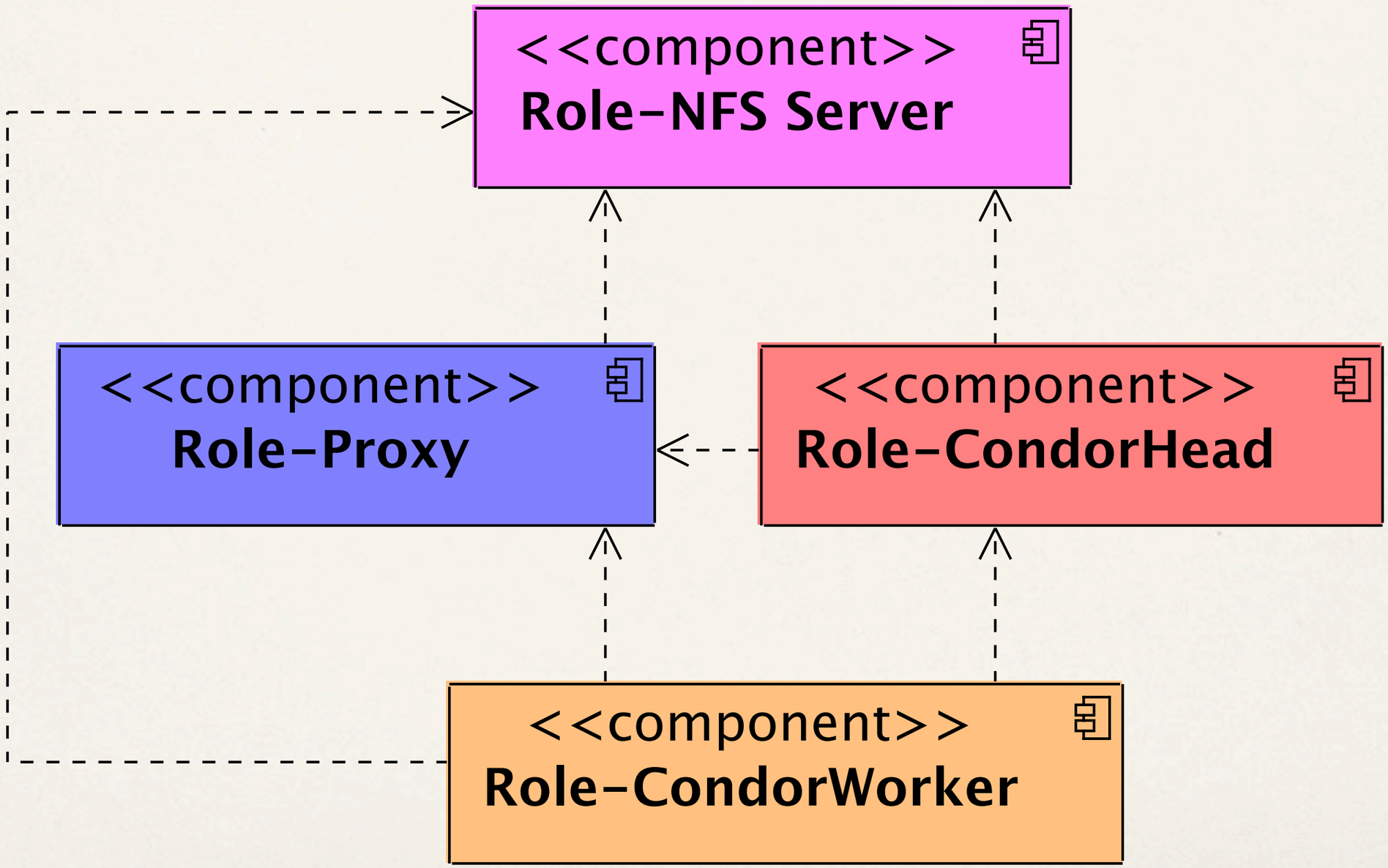


Design a cluster in days
e.g. the tier3 workgroup will design a tier3 cluster and suggest it to the sites

Setup and deploy a cluster in hours onto a physical or Virtual Machines

Feels no difference

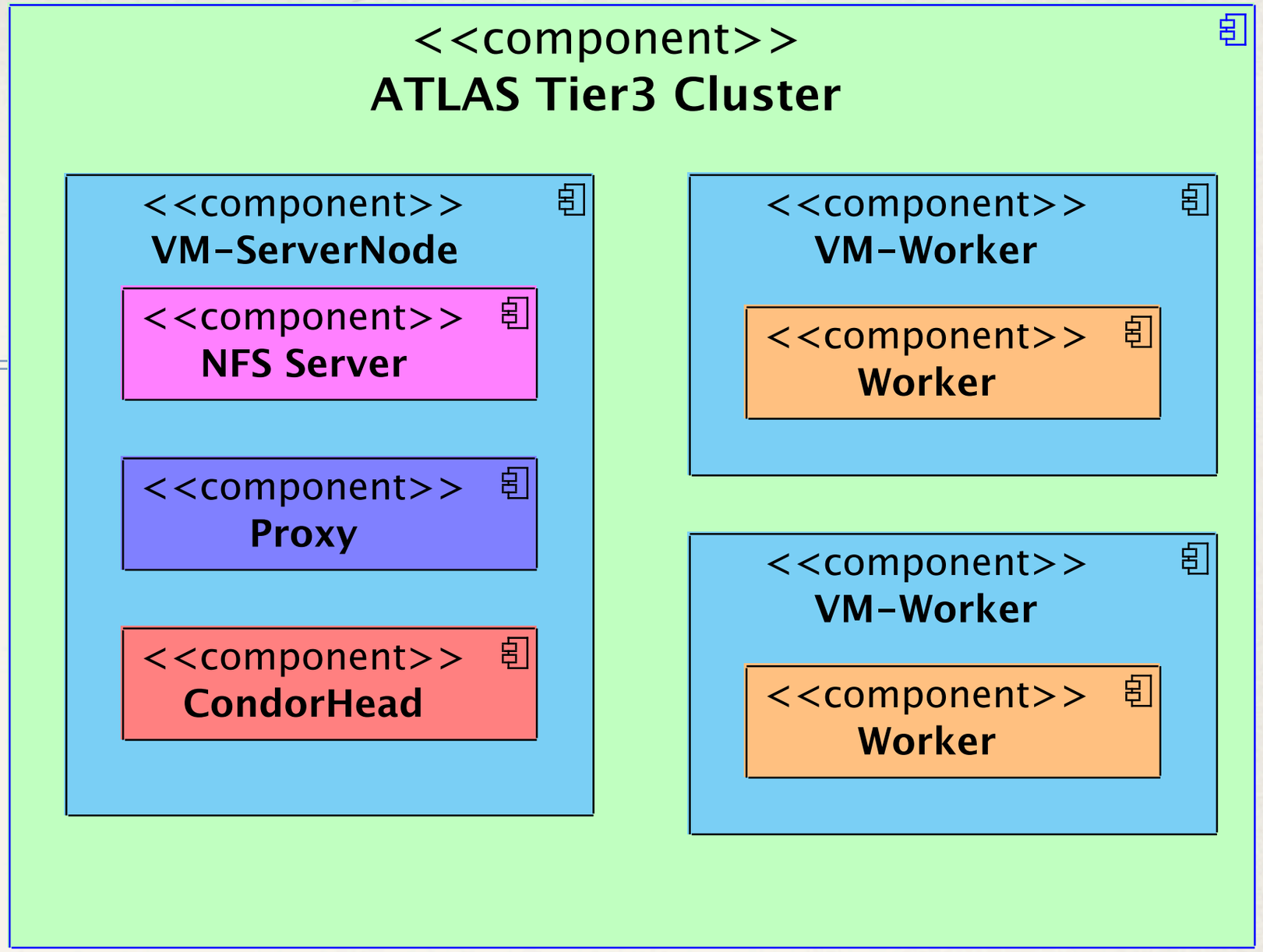
A Simplified Tier3 Cluster and Its Roles

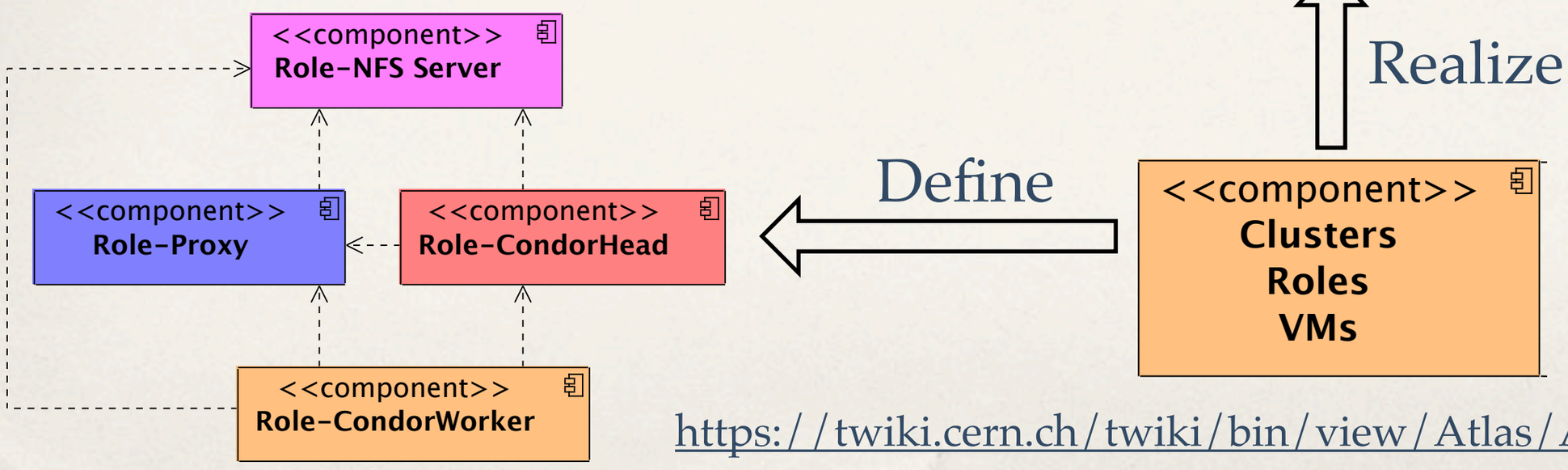


- "Cluster-Role-VM" is a tool to automatically "deploy(contexualize)" a defined set of roles on a phy-cluster/cloud.

- Each role will have a status "On/Off" that can be dynamically controlled from outside (e.g. scheduler)

- CRV will make sure all "on" roles are running. And "off" roles stopped.





https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasTier3VMCRV

Scheduler (e.g. Condor) feedback

<<component>> Puppet

<<component>> BCFG2

<<component>> IProfileMaker

Apply Roles

<<component>> Clusters Roles VMs

REST

VM

VM

VM

Web-Interface Control, monitor

<<component>> IVMController

Start/Stop

<<component>> Eucalyptus/EC2

<<component>> Nimbus

<<component>> Physical Machines

The status of the Cluster will be monitored and can be changed